UNIVERSITY OF LIVERPOOL

Newcastle University

Centre for Data

ESRC Smart Data Research
UK Strategic Advice Team

REPORT 1: SAT ACTIVITIES DATA SERVICES & THEMATIC RESEARCH PILLARS

Rachel Franklin, Simeon Yates,
Elena Musi, Omar Guerrero, Seth Spielman,
Jeanette D'Arcy, Jessica Crosby

April 2023

DMSI
Digital Media and Society Institute

Centre for Urban & Regional Development Studies
CURDS

CONTENTS

INTRODUCTION

Please note: This preliminary report was produced in April 2023. Since the writing of this document, the name of the programme and its attendant terminology has changed from 'Digital Footprints' (DF) to 'Smart Data' (SD).

This report was commissioned by Smart Data Research UK – an Economic and Social Research Council data infrastructure programme. The purpose of the report is to provide context, input and early-stage recommendations regarding the overall structure of the Digital Footprints Phase 2 Investment. It was prepared by Jeanette D'Arcy at the University of Liverpool. This researcher is part of a group established by Smart Data Research UK to give independent strategic advice between 2022 and 2024.

Disclaimer

The findings and conclusions presented in this report are solely those of the researchers and do not necessarily reflect the views of Smart Data Research UK or the Economic and Social Research Council.

With the Digital Footprints (DF) programme, the ESRC has enormous potential to generate a step change in the use of Digital Footprints Data (DFD) for novel and impactful social and economic research that will improve lives, contribute to economic development, and provide valuable insights into the world we inhabit. The DF investment comes at a time of increased recognition of the power of data, data science and data infrastructure for driving social and economic innovation, including the UK's National AI Strategy (Her Majesty's Government, 2021), UKRI's Digital Research Infrastructure Programme (DRI), the Data and Analytics Facility for National Infrastructure (DAFNI), the Alan Turing Institute's transition to 'Turing 2.0' and NERC's investment in digital solutions and data infrastructure. Similar institutional discussions are underway elsewhere, including in the United States, with the National Academies of Sciences, Engineering, and Medicine's Committee, "Toward a Vision for a New Data Infrastructure for Federal Statistics and Social and Economic Research in the 21st Century", as well as Australia's Research Data Commons. Linking initiatives, building connections, promoting programme visibility, and providing evidence for best practices has never been more important.

Our impetus for undertaking this work and convening this Strategic Advice Team (SAT) lies in our belief that coordinated efforts around DFD can revolutionise social science understanding of the grand challenges that face our world and our capacity to address them—but only with intentional efforts to understand and shape the existing DFD landscape. Over the next two years, the SAT will consider the key challenges around data access, infrastructure, integration, and capacity, but by design also leave room for additional perspectives to emerge throughout the range of planned engagement activities. Digital data—the various digital footprints and traces left by citizens and organisations—plays a key role in economic innovation and growth, government administration and, of course, research. National research councils hold enormous sway in helping to make this data accessible to researchers and citizens alike, but also play a unique role: the *capacity* to strategically invest and the *visibility* to set expectations around data access, integration, and utility, as well as training, infrastructure, ethics, and capacity building.

This preliminary report provides context, input and early-stage recommendations regarding the overall structure of the Digital Footprints Phase 2 Investment, with particular focus on Thematic Research Pillars and data services. These findings are the result of conversations with the Strategic Advice Team (SAT) Advisory Group, as well as two online workshops conducted in May and April 2023. Further details are provided below.

EXECUTIVE SUMMARY AND HEADLINE RECOMMENDATIONS

Expanded discussion of recommendations and findings can be found in Sections 4 and 5. Below we summarise our main preliminary recommendations:

1. There is **widespread support** in the community for the Digital Footprints investment and documented need for the data the investment seeks to provide. Data Services should help to reduce 'red tape', facilitating safe and secure access to data and tackling barriers to access, especially with regard to private sector and/or sensitive data.
2. Two possible frameworks for **design of Data Services** emerged: 1) a pragmatic approach based on where/from whom data will be sourced, what data is available, data types and markets; 2) an open approach based on the objectives of social science research, methodologies, and themes.
3. Data Services should not be siloed, separate entities. Researchers are very likely to want access to more than one Service's data and to this end, **interoperability and linkage** must be a 'baked in' consideration from the beginning, with the goal of balancing the needs of those running the Services and those wanting to use the data for research. This has important implications for research infrastructure and access.
4. Although this is primarily an infrastructure investment, **skilled support and/or training** for researchers is a necessity in order for researchers to access and work with complex, integrated data sets and decentralised analytics.
5. There remains some **confusion** in the community about what constitutes "Digital Footprints" data. Clarification will be important in order for programme investments to be efficiently run and successful.

METHODOLOGY AND WORK COMPLETED

To date, we have conducted two online sessions with experts in the Advisory Group (see Appendix four) centred around launching the project and responding to and developing findings from Workshop 1. Each session lasted one hour. For the second session, Advisory Group members were provided a copy of the ESRC's envisioned DF programme structure (Fig.1) and given a series of prompts based on this and the discussion points from Workshop 1. Members of the SAT were in attendance to listen, prompt, make notes, discuss and summarise the outcomes of this engagement, the results of which are presented in the findings section below.

With the support of The Collective as facilitators, we have delivered two three-hour Engagement and Consultation workshops targeting different stakeholders. These workshops are designed to build collaborative interactions which outlive the network and help develop a broader research community (Bramley and Ogilvie, 2021). Each workshop reflects on the outcomes from prior workshop rounds and participants are encouraged to attend more than one workshop.

Open invitations were distributed via relevant academic mailing lists, as well as through the team's own networks, and participants were invited to register via Eventbrite. The first workshop, 'Research and Data Priorities' (Workshop 1), took place on 7th March 2023 with more than forty in attendance from a range of academic disciplines. The second, 'Identifying Researcher Needs (Workshop 2), took place on 18th April 2023 with 34 attending, also from a range of academic disciplines (see Appendices for workshop questions and range of disciplines). The data gathered from participants in these workshops are presented in the findings section below. Participants were first encouraged to 'break the ice' with a short introductory group session, then divided into Breakout room discussion groups which were curated to have a mixture of both academic disciplines and career levels in each. Groups were given a discussion prompt (see Appendices 1 and 2) and asked to both feed back orally at the end of the discussion, and to make notes on a provided Google Doc. Again, members of the SAT were in attendance to listen, prompt, make notes and summarise the outcomes of these discussions.

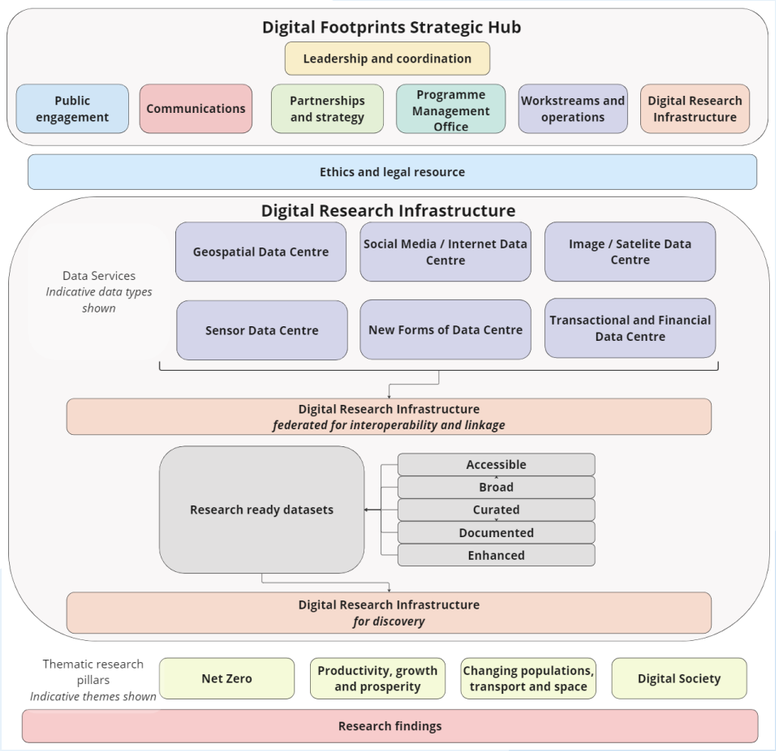

Fig.1. *Envisioned DF programme structure, ESRC, as of November 2022. Please note this has been refined and revised since the writing of this document*

FINDINGS

This section provides preliminary findings in relation to the aims of the programme: to identify and document the potential DFD offers; to identify and document existing barriers to fulfilling that potential; and to provide informed advice to ESRC, especially with regard to Thematic Pillars and Data Services.

The overarching themes that emerged from analysis of Workshop 1, Workshop 2 and Advisory Group sessions include: 1) Benefits of access to DFD, 2) Barriers to access to DFD, 3) Thematic Pillars, 4) Data Service categorisation, 5) Interoperability, linking and integration, 6) Ethics and legalities.

Benefits of access to DFD

Researchers were enthusiastic about the area and the workshops demonstrated the broad range of research that is currently being done in this area, as well as what data researchers would like to have access to in the future. Some of the types of DF data and work mentioned in 'wish lists' are listed below. Inclusion of some non-DFD data types may reflect community fuzziness around what constitutes "digital footprints" data.

- Mobile phone app data
- Health surveillance, health records on specific conditions, emergency service records
- Traditional data (e.g., annual surveys)
- Remote imagery data
- Crime/court records
- Visual imagery on urban areas
- Labour market and economy, incomes, industries, occupations
- Co2 emissions
- Property data, smart metre data
- Social media data
- Human mobility data, geospatial data
- Decisions made by algorithms about people
- Social capital, place attachment, relational measures to *'balance out more qualitative data'* (Workshop 2)

Workshop discussion groups spoke about beneficial social factors of working with DFD, especially with respect to (mental) health outcomes (Workshop 1), making data consultation more inclusive and increasing public participation (Workshop 1). Groups pointed out the beneficial possibility of focusing on particular types of applications, such as longitudinal research or linked data; pointing out that a particular advantage of DFD is being able to trace people over time and space; to measure things in real time (Workshop 1). While researchers felt that work with DFD can have great social impact, they also discussed many barriers to discovering, accessing and using DFD.

Barriers to access to DFD

In terms of discoverability, participants questioned if researchers know, or can find out, what data is available, and pointed out that there could be lot of duplication of work when researchers individually contact companies/providers to ask if/what data is available (Workshop 2). Once the data has been discovered, there is often a lack of structure in available data sets, it is often not ready for analysis in its raw form (Workshop 1), can be messy or incomplete, and researchers can spend valuable time cleaning data only to end up with something that is not as suitable as they originally thought (Workshop 2).

Availability of data can also be uncertain as data owners may change what is accessible, e.g., platforms which allow users to scrape data may change this policy and researchers then have to start from scratch (Workshop 2). Private sector data, especially sensitive and confidential data, is difficult to get access to and highly sought after (Workshop 2). Researchers can encounter resistance to using data, due to a variety of reasons including fear of the unknown, as data owners may not understand the potential benefits of sharing and may not want to risk sharing when they are uncertain of outcomes or risks (e.g., to reputation) (Workshop 2). The format of data and frequency with which it can be accessed is also dependent on how data owners make their data available, and what type(s) of data they hold. Legal and ethical issues are also sometimes unclear and subject to change; evolving terms and conditions create instability and raise questions about replication of studies in the future (Workshop 2). Several groups raised the challenge of data contracts, which can be difficult to obtain and costly in terms of time commitment, especially under the current system of short-term grants leading to time constraints. Consequent pressures were put into perspective by an example where a researcher had two years to complete a project but it took 18 months to secure a contract to get access to the necessary data (Workshop 2).

Once the availability of data has been established and researchers gain access, analysis can also be challenging. Participants raised the question that, if researchers are using new methods, what theoretical underpinning are social science researchers drawing on? Discussions pointed out that traditional methods for knowledge generation in science can be difficult to apply to DFD, and findings can be speculative (Workshop 2). Working across geographical boundaries can make research difficult to navigate as researchers must operate under different legal frameworks. Workshop participants suggested it could be useful to have a 'checklist' of all the elements that researchers must consider

in order to understand the bigger picture of legalities, ethics, ownership, etc (Workshop 2).

Thematic pillars

Suggestions for pillars were largely in line with those proposed by ESRC in their envisioned Digital Footprints programme structure (Fig. 1). However, some notable exceptions were:

i) An emphasis placed on ethics and inequalities, not just in terms of how the programme itself will treat ethical considerations and legalities, or how the programme will address questions of equity, but in terms of research outputs;

ii) A proposed thematic grouping around key 'challenges';

iii) The consideration of citizen-led/open data principles.

Some specific ideas emerging from Workshops for potential pillars were:

- Health, population size, inequalities, inequalities of access
- Access to technology
- Access and accessibility
- Link to SDGs
- Organise pillars around 'challenges' associated with DFD
- Social connectedness and social Engagement
- Ethics of DFD
- Under-represented groups
- Wellbeing in the digital environment; social justice
- Ethics (inclusivity and equity)
- Environment (esp. climate change) and society
- Security and online safety
- Inequalities; policymaking; North-South divide and levelling up;
- Sustainability; economy and employment
- Mobilities

There was curiosity expressed about the differences between 'pillars' and 'services' and how these might link and/or interact.

Data centre categorisation

The possibilities for ways of organising or classifying data centres was returned to several times in discussion and there were several useful suggestions both from Workshop 2 and the Advisory Group engagement session. There was considerable overlap between suggestions for thematic pillars and for data services centres, which included:

- **Organisation around data types**, in a similar structure to that currently envisioned by the ESRC (Fig. 1). This suggests a pragmatic approach, based on where/from whom data will be sourced, and what data is available. Participants acknowledged this was a logical way to structure data services, but there were questions over how this would relate to themes, whether there would be enough interoperability and concerns that this approach could end up in siloed services focused on their own interests.
- **Organisation around themes**, for example economic activity, health or public discourse. This was felt to be an approach that would be more likely to appeal to, and be driven by, social scientists, but that it would also face issues around interoperability and integration as researchers would want to be able to access data relating to more than one theme.
- **Organisation around markets or industries**, for example population movement data, transport data, property data; this would enable the approach of bringing together providers who work with this type of data with a team to focus on procurement and commercial expertise. This would still need a focus on bringing the data together across different services but would encourage a common methodology for accessing and storing data.
- **Organisation around methods**, for example: visual (a distinction made between 'image' here, as 'visual' refers to method of analysis), predictive, spatial. This has the advantage of operating across interpretive and quantitative methods and would encourage a focus on the endpoint of the research to be produced as outputs. Similarly, workshop discussions raised the possibility of focusing on particular types of application, such as longitudinal research or linked data.
- Some specific ideas for potential data services centres were i) longitudinal studies, ii) linked administrative data, iii) longitudinal population studies, iv) smart cities, v) data governance and ownership (Note: in the main these are not Digital Footprints data; we include for context).

It was suggested that discussions which focus on research outputs ('substantive questions') help in thinking about how to integrate data and are more useful for social science. **Thinking about how to integrate data without considering research questions could end up with data sets that will not be used.** By starting with an idea of what outputs researchers would want to produce, the necessary infrastructures could be 'reverse engineered'. Some specific outputs brought up in discussions were:

- *'Using DFD to test/identify relationships between features of these footprints and patterns of online victimization'* (Workshop 2)
- *'Using DFD to analyse the personal narratives available online and comparing AI and human narratives; bridging the divide between big data and the personal'* (Workshop 2)
- *'Understanding the main types of planning activities in the UK and how improving the planning system through digitalisation can make the process better, in combination, e.g., with housing demand/housing stock information*

A technical coordinating body could enable a conversation between the objectives of social science research and the practical considerations of data integration.

Workshop discussions pointed out that it will be important to build on existing services but to fill any gaps, rather than replicate what is already being done. Workshop groups suggested taking into consideration how existing centres have developed to inform future development (Workshop 1), looking at past failures to make the best of learning from what has previously been done (Workshop 2) and thinking of infrastructure as long-term to avoid loss of data and expertise (Workshop 1).

Workshop participants discussed the potential 'mechanisms' that could be put in place between researchers and providers to help access particular forms of data that require partnership or legal agreements, suggesting that 'data mediators' could help researchers overcome many issues (Workshop 2).

It is worth noting that several workshop groups indicated that the format in which data is made available is not necessarily important, and it was generally agreed that synthetic data would be extremely useful as well as allowing researchers to avoid many issues. A suggested way to provide access was to have downloadable 'dumps' of data to use short-term APIs to build specific data sets that suit researchers' needs.

Interoperability, linking and integration

From the discussions at both Workshops 1 and 2, and expert opinion from the Advisory Group, linking and integration make up a crucial component of DF programme success. Experts explained that social science researchers will want access to more than one data type. Therefore, especially if the data services are categorised by data type, this will require interoperability to be 'baked-in' from the initial set-up, along with evolving considerations of ethics and legalities. Workshop 1 discussions raised concerns about whether domain-areas could cause DFD funding to be siloed or limit *collaboration across disciplines.* (Workshop 1). The importance of **data integration and combining DFD with traditional forms of data** came up several times, along with considerations of how the DF programme might connect to others such as Administrative Data Research (ADR) UK. For experts, interoperability was felt to be important not just for the programme now, but for future directions of research, with the example raised of 'citizen social science': the methodologies of this work do not fit well with the model of categorisation by data type and would require careful consideration of how to curate data. It was suggested that 'citizen data' could be another named data service , as this would encourage the working through of the above considerations. Workshop discussions encouraged following Open Science principles. When asked about infrastructure, Workshop 2 participants responded that ideally, basic interoperability should be pre-created.

A highlight from both Workshop 1 and the Advisory Group engagement sessions was the potential for following the example of Trusted Research Environments (TREs) and how these are currently working in, e.g., healthcare research (see Varma et al, 2021, DARE UK 2023). It was acknowledged that analysing federated data can be challenging and data services will need to work with researchers to carry this out. An important consideration here would be that data services would need to have local processing power rather than storage only, to facilitate analysis.

It was understood that such goals of interoperability are difficult and long-term; it may be that while linking and integration are a key part of the design of the programme, the achievement of such linking would be a long-term goal rather than a short-term one. Similarly, Workshop 1 participants acknowledged more generally that the kind of work being planned in this programme will not be achieved quickly, that it is important to consider time horizons for investments as not only do researchers need to know that data will be available when they need them, but investments also need time to develop relationships with data owners. Ethical issues were also raised, as linking data to other data sets means it cannot be fully anonymised, with implications for security and privacy.

Ethics and legalities

Many participants brought up the importance of ethical considerations. Several groups spoke about issues of safety, privacy and transparency. One group voiced concerns particularly around the ethics of linking data but pointed out that this was balanced against value to be gained in linking data, e.g., in health and education. One group asked how it will be possible to conduct longitudinal data collection in ways that are both private and secure, and meaningful (Workshop 1).

Many discussed the impact of DFD on social inequalities and divisions. One group asked which groups have their footprints collected, and what assurances there are that aggregation will lead to good outcomes. This is also important in terms of the social science objectives of the DF programme, as researchers spoke about the importance of using DFD to discover inequalities and/or detect discrimination. They pointed out that available data sets are often not fit for purpose in this respect as they may not include all the necessary metadata (Workshop 2).

Several discussed relationships between private and public sector. One group gave an example of Airbnb data, which is legal to scrape but then cannot be shared, and asked how a legal framework could overcome situations such as this one. Several groups noted the issue of whether or not, and to what extent, people know they are sharing their data and how this can then be further shared ethically. Workshop participants pointed out that systems which store a copy of data in a 'safe haven' are reassuring for both regulators and researchers (Workshop 2). Workshop participants discussed inequalities in access inherent when availability of particular forms of DFD is reliant upon leveraging relationships with data owners/ partners/providers (Workshop 2).

An important point raised was the need for agreed standards, protocols and procedures, e.g., how to assess the quality of DFD such as social media data. Metadata was raised as a 'hot topic' (Workshop 2), with discussions asking how to evaluate and standardise descriptions of how data was created and why data is labelled in a particular way (Workshop 2). Metadata was discussed as part of a wider approach focused on 'open processes', as it can provide a richer description of a better sense of how data is generated from end to end and how it has been shaped by those processes. (Workshop 2). It was pointed out that data accessed by researchers is often provided without metadata on how it was gathered, what is included or excluded and why (Workshop 2).

PRELIMINARY RECOMMENDATIONS

Thematic Pillars and Data Service Centres

While discussions were largely in agreement with the current organisation of thematic pillars, there were important suggestions for alternate arrangements of data service centres. Overall, data services should help to reduce 'red tape', facilitating safe and secure access to data and tackling barriers to access, especially with regard to private sector and/or sensitive data.

At this stage in the project, no one clear strategy for the organisation of data service centres has emerged, but findings fall broadly under two possible approaches:

1. A pragmatic approach based on where/from whom data will be sourced, what data is available, data types and markets.
2. An approach based on the objectives of social science research, methodologies, and themes.

These two approaches may yield very different types of proposals. The first approach creates some efficiencies in data collection, as a more easily identifiable and narrow range of providers would be targeted. It has the disadvantage of not aligning with specific research questions, or the expertise of researchers. The second approach, more driven by research questions, may be more attractive to researchers though less tractable from the perspective of data gathering. The second approach also would require more coordination among teams to prevent overlapping collections. As the project develops, this will be an important area for interviews and survey questions to focus on.

- Options:
  - Organisation of thematic pillars/data services grouped around key 'challenges'
  - Organisation of data services around data types (this could ease development and management of relationships with data providers).
  - Addition of data services focused on best practice mechanisms for data linkage, citizen/civic data, and ethics/legality (see below)
  - Use and creation of synthetic data by services for exploration, testing and development
  - Creation of a technical coordinating centre/body to address the practical concerns of data integration and linkage to address the objectives of social science research
  - Organisation of services around specific methods rather than data type – or a linkage of centres with methods foci

The workshops and board sessions raised the importance of linking data types/centres with methods training. As noted below in 5.3 there is a need to link the curation of data with providing social researchers with the necessary methodological skillsets, especially if a key goal of the programme is developing the UK Digital Footprints infrastructure, skillsets, and international leadership.

Interoperability, linking and integration

A key concern raised was that data service centres should not be siloed, separate entities. Researchers are very likely to want access to more than one centre's data. To this end, interoperability must be a 'baked in' consideration from the beginning, with the goal of balancing the needs of those running the service centres and those wanting to use the data for research. While linking data was a key concern for researchers and experts that was raised many times, there was no one clear solution as to how this should be done, whether through curated data sets, or federated. As the project develops, this will be an important focus, for example upcoming interviews with data owners will be used to shed light on possible approaches. Some researchers will want data that is already linked, others will want to integrate data more flexibly.

- Options:
  - Creation of a data service whose focus is on best practice for data linkage, both providing and storing data in a manner that allows it to be linked and providing advice on the technical aspects of how this linking can be achieved. At the extreme, individual data centres might outsource the storage and distribution of data to a central "hub" organization as this might reduce the costs/complexity of building multiple similar data infrastructures. Alternatively, the hub might provide a reusable technical template to organizations
  - Creation of federated platforms, or a central shared 'hub' platform for sharing data, collaboration and support
  - Services doing the work of providing common keys that would allow researchers to join up data
  - Creation of 'safe havens' for processing and analysis, following the examples of Trusted Research Environments. It is important, however, that these are accessible remotely rather than requiring travel to physical locations, and that service centres are equipped with enough processing power for analysis as well as storage
  - Consideration of open research and citizen/peoples' data models, including the creation of a data service centre focused on citizen access and use

Skillsets and recruitment

To access and work with complex, integrated data sets and decentralised analytics, skilled support and/or training for researchers is a necessity. Researchers do not necessarily have the skills to be able to access data in the ways they would like to, and the subject of skills was enthusiastically raised many times. While the main focus of this programme is infrastructure, arguably the infrastructure will not be useful without the skills necessary to access it.

- Options:
  - Requirement in calls for bidding universities to provide clear plans for career support
  - Provision of training and support for data science and programming skills, including provision for high-level, long-term recruitment

Ethics and legalities

The difficulties and uncertainties of obtaining licensing agreements, and ethical considerations of both DFD use and the study of how DFD affect ethics and inequalities are key concerns.

- Options:
  - Provision of expert advice and support in terms of legal and ethical issues, including a 'checklist' document for researchers of elements that must be considered for all projects in terms of legalities, ethics, ownership, etc
  - Development of standardised documentation (e.g., data agreements) to be used across Data Services (and possibly beyond)
  - Consideration of how to lead by example in terms of equitable access, metadata management, and other ethical concerns
  - Development of procedures to ensure that information about individuals and sensitive groups is protected from disclosure

NEXT STEPS

1. Continued facilitation of workshop series. ‘Learning from International Best Practice’, is planned for 18th May 2023, ‘Infrastructure’ (title tbc) for the 7th June 2023. There will be a ‘Meet the ESRC DF Team’ event on 27th June.
2. Interviews with industry stakeholders (i.e., data owners) and the Advisory Group. These will be arranged and conducted over May/June/July 2023. It was decided that a one-to-one approach would be more appropriate for industry participants than a workshop environment. Stakeholders will be identified via the ESRC and SAT’s contacts and invited for interview, with snowballing then used to identify further participants.
3. Administration of a Delphi review survey to seek consensus about the options and requirements for a successful DF programme.

REFERENCES

Bramley, A., & Ogilvie, L. (2021). Research Collaboration: A step-by-step guide to success. IOP Publishing.

HM Government. (2021). National AI Strategy. National AI Strategy - GOV.UK (www.gov.uk)

UK Health Data Research Alliance, & NHSX. (2021). Building Trusted Research Environments - Principles and Best Practices; Towards TRE ecosystems (1.0). Zenodo. https://doi.org/10.5281/zenodo.5767586

DARE UK Delivery Team. (2023). Federated Architecture Blueprint. Federated Architecture Blueprint (dareuk.org.uk)

APPENDIX ONE

ESRC Digital Footprints: Workshop 2

Research and Data Priorities

Questions for Breakout Room Discussions

BREAKOUT DISCUSSION 1:

What does digital footprint data mean to you in social science research?

BREAKOUT DISCUSSION 2

What might the thematic research areas be where digital footprint data might make the biggest difference?

BREAKOUT DISCUSSION 3
How might the data centres be organised?

Should these be based on Research Topic/Data type/Method/ Discipline?

APPENDIX TWO

ESRC Digital Footprints: Workshop 2

Identifying Researcher Needs

Questions for Breakout Room Discussions

BREAKOUT DISCUSSION 1:

Reflecting on your experience, what research would you like to do with digital footprints data but currently can't ?

Please highlight any issues of equity and use.

BREAKOUT DISCUSSION 2

Thinking of your wishlist -what digital footprints data would you like to have access to?

- With what characteristics?
- In what format?
- At what frequency- how often?

BREAKOUT DISCUSSION 3

Thinking of our last discussion:

- What infrastructure would the ESRC need to support to make this happen?
- How would you like to access the digital footprints data within this infrastructure?

APPENDIX THREE

Academic Disciplines: Workshop 2

To ensure diversity in discussion groups, attendees were asked their academic discipline and grouped accordingly. A wide variety of academic disciplines were represented. The following disciplinary groups were identified:

ACCOUNTING

ANTHROPOLOGY

ARCHIVAL SCIENCE

BEHAVIOURAL SCIENCE

BIG DATA ANALYTICS

COMMUNICATION AND MEDIA

COMPUTATIONAL SOCIAL SCIENCE

DATA SCIENCE/DATA ENGINEERING

DEMOGRAPHY

DESIGN

DIGITAL HEALTH/MEDICINE

EDUCATION

ENERGY AND ENVIRONMENT

ENGLISH LITERATURE

FORCED MIGRATION

GEOGRAPHY/HUMAN GEOGRAPHY/ECONOMIC GEOGRAPHY/GIS

GEOSPATIAL SCIENCE

HERITAGE/DIGITAL CULTURE

LAW

LINGUISTICS/SPEECH TECHNOLOGY

MANAGEMENT SCIENCE

PARASITOLOGY

PSYCHOLOGY

SOCIOLOGY

URBAN PLANNING/URBAN STUDIES/REAL ESTATE FINANCE

DMSI
Digital Media and Society Institute

Department of Communication & Media
University of Liverpool
School of the Arts
19 Abercromby Square
Liverpool
L69 7ZG

Newcastle University
Centre for Data

Centre for Urban & Regional Development Studies
CURDS

THE ORIGINAL REDBRICK